Bernoulli **17**(1), 2011, 424–440
DOI: 10.3150/10-BEJ276

On the functional central limit theorem via martingale approximation

MIKHAIL GORDIN[1] and MAGDA PELIGRAD[2]

[1] *POMI (Saint Petersburg Department of the Steklov Institute of Mathematics), 27 Fontanka emb., Saint Petersburg 191023, Russia. E-mail: gordin@pdmi.ras.ru*
[2] *Department of Mathematical Sciences, University of Cincinnati, P.O. Box 210025, Cincinnati, OH 45221-0025, USA. E-mail: peligrm@ucmail.uc.edu*

In this paper, we develop necessary and sufficient conditions for the validity of a martingale approximation for the partial sums of a stationary process in terms of the maximum of consecutive errors. Such an approximation is useful for transferring the conditional functional central limit theorem from the martingale to the original process. The condition found is simple and well adapted to a variety of examples, leading to a better understanding of the structure of several stochastic processes and their asymptotic behaviors. The approximation brings together many disparate examples in probability theory. It is valid for classes of variables defined by familiar projection conditions such as the Maxwell–Woodroofe condition, various classes of mixing processes, including the large class of strongly mixing processes, and for additive functionals of Markov chains with normal or symmetric Markov operators.

Keywords: conditional functional central limit theorem; martingale approximation; mixing sequences; reversible Markov chain

1. Introduction and results

The objective of this paper is to find a characterization of stationary stochastic processes that can be studied via a martingale approximation in order to derive the functional central limit theorem for processes associated with partial sums.

There are several ways to present the results since stationary processes can be introduced in several equivalent ways. We assume that $(\xi_n)_{n\in\mathbb{Z}}$ denotes a stationary Markov chain defined on a probability space $(\Omega,\mathcal{F},P)$ with values in a measurable space $(S,\mathcal{A})$. The marginal distribution and the transition kernel are denoted by $\pi(A)=P(\xi_0\in A)$ and $Q(\xi_0,A)=P(\xi_1\in A|\xi_0)$, respectively. In addition, Q denotes the operator acting via $(Qf)(\xi)=\int_S f(s)Q(\xi,\mathrm{d}s)$. Next, let $\mathbb{L}_0^2(\pi)$ be the set of functions on S such that $\int f^2\,\mathrm{d}\pi<\infty$ and $\int f\,\mathrm{d}\pi=0$. Denote by $\mathcal{F}_k$ the σ-field generated by ξ_i with $i\le k$, $X_i=f(\xi_i)$, $S_n=\sum_{i=0}^{n-1}X_i$ (i.e., $S_0=0,S_1=X_0$, $S_2=X_0+X_1,\ldots$). For any integrable

This is an electronic reprint of the original article published by the ISI/BS in *Bernoulli*, 2011, Vol. 17, No. 1, 424–440. This reprint differs from the original in pagination and typographic detail.

variable X, we define $\mathbb{E}_k(X) = \mathbb{E}(X|\mathcal{F}_k)$. In our notation, $\mathbb{E}_0(X_1) = Qf(\xi_0) = \mathbb{E}(X_1|\xi_0)$. We also set $\mathcal{F}_{-\infty} = \bigcap_{k\in\mathbb{Z}} \mathcal{F}_k$.

Throughout the paper, we assume $f \in \mathbb{L}_0^2(\pi)$; in other words, we assume that $\|X\|_2 = (\mathbb{E}[X_1^2])^{1/2} < \infty$ and $\mathbb{E}[X_1] = 0$.

Note that any stationary sequence $(Y_k)_{k\in\mathbb{Z}}$ can be viewed as a function of a Markov process $\xi_k = (Y_i; i \le k)$ for the function $g(\xi_k) = Y_k$.

The stationary stochastic processes may be also introduced in the following, alternative, way. Let $T : \Omega \mapsto \Omega$ be a bijective bimeasurable transformation preserving the probability. Let $\mathcal{F}_0$ be a sub-σ-algebra of $\mathcal{F}$ satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$. We then define the non-decreasing filtration $(\mathcal{F}_i)_{i\in\mathbb{Z}}$ by $\mathcal{F}_i = T^{-i}(\mathcal{F}_0)$. Let X_0 be a random variable which is $\mathcal{F}_0$-measurable. We define the stationary sequence $(X_i)_{i\in\mathbb{Z}}$ by $X_i = X_0 \circ T^i$.

In this paper, we shall use both frameworks.

In order to analyze the asymptotic behavior of the partial sums $S_n = \sum_{i=0}^{n-1} X_i$, Gordin, in [15], proposed to decompose the sums related to the original stationary sequence into the sum

$$S_n = M_n + R_n \tag{1}$$

of a square-integrable martingale $M_n = \sum_{i=0}^{n-1} D_i$ adapted to $\mathcal{F}_n$, whose martingale differences (D_i) are stationary, and a so-called coboundary R_n, that is, a telescoping sum of random variables with the basic property that $\sup_n \mathbb{E}(R_n^2) < \infty$. More precisely, $X_n = D_n + Z_n - Z_{n-1}$, where Z_n is another stationary sequence in $\mathbb{L}_2$. The limiting properties of the martingales can then be transported from the martingale to the general sequence. In the context of Markov chains, the existence of such a decomposition is equivalent to the solvability of the Poisson equation in $\mathbb{L}_2$.

For proving a central limit theorem for stationary sequences, a weaker form of martingale approximation has been pointed out by many authors (see, e.g., [21] for a survey). Recently, two interesting papers, one by Dedecker, Merlevède and Volný [7] and the other by Zhao and Woodroofe [32], provided necessary and sufficient conditions for martingale approximation with an error term in (1) satisfying

$$\mathbb{E}((S_n - M_n)^2)/n \to 0. \tag{2}$$

This decomposition is strong enough for transporting the conditional central limit theorem from sums of stationary martingale differences in $\mathbb{L}_2$ to $S_n/\sqrt{n}$. By conditional CLT, as discussed in [6], we understand, in this context, that for any continuous function f such that $|f(x)|/(1 + x^2)$ is bounded and for any $k \ge 0$,

$$\left\| \mathbb{E}_k(f(S_n/\sqrt{n})) - \int_{-\infty}^{\infty} f(x\sqrt{\eta})g(x)\,\mathrm{d}x \right\|_1 \underset{n\to\infty}{\longrightarrow} 0, \tag{3}$$

where g is the standard normal density and $\eta \ge 0$ is an invariant function satisfying

$$\lim_{n\to\infty} \left\| \frac{\mathbb{E}_0(S_n^2)}{n} - \eta \right\|_1 = 0.$$

Here, and throughout the paper, we denote by $\|\cdot\|_p$ the norm in $\mathbb{L}_p$.

An important extension of this theory is to consider the conditional central limit theorem in its functional form. For $t \in [0,1]$, define

$$S_n(t) = S_{[nt]} + (nt - [nt])X_{[nt]},$$

where $[x]$ denotes the integer part of x. Note that $S_n(\cdot)/\sqrt{n}$ is a random element of the space $C([0,1])$ endowed with the supremum norm $\|\cdot\|_\infty$. Then, by the conditional CLT in the functional form (FCLT), we understand that for any continuous function $f : C([0,1]) \to \mathbb{R}$ such that $x \mapsto |f(x)|/(1+\|x\|_\infty^2)$ is bounded and for any $k \geq 0$, we have

$$\left\| \mathbb{E}_k(f(S_n/\sqrt{n})) - \int_{C([0,1])} (f(x\sqrt{\eta}))\, \mathrm{d}W(x) \right\|_1 \underset{n\to\infty}{\longrightarrow} 0. \tag{4}$$

Here, W is the standard Wiener measure on $C([0,1])$.

It is well known that a martingale with stationary differences in $\mathbb{L}_2$ satisfies this type of behavior with $\eta = \lim_{n\to\infty} \sum_{l=0}^{n-1} D_l^2/n$ in $\mathbb{L}_1$ – this is at the heart of many statistical procedures. This conditional form of the invariance principle is a stable type of convergence that makes possible the change of measure with another absolutely continuous measure, as discussed in [1, 11, 27].

With such a result in mind, the question is now to find necessary and sufficient conditions for a martingale decomposition with the error term satisfying

$$\mathbb{E}\Big(\max_{1\leq j\leq n} (S_j - M_j)^2\Big)/n \to 0. \tag{5}$$

In order to state our martingale approximation result, for fixed m, we consider the stationary sequence

$$Y_0^m = \frac{1}{m}\mathbb{E}_0(X_1 + \cdots + X_m), \qquad Y_k^m = Y_0^m \circ T^k. \tag{6}$$

In the language of Markov operators, we then have

$$Y_0^m = \frac{1}{m}(Qf + \cdots + Q^m f)(\xi_0).$$

It is convenient to introduce a seminorm notation, namely,

$$\|Z\|_{M^+} = \lim\sup_{n\to\infty} \frac{1}{\sqrt{n}} \left\| \max_{1\leq k\leq n} \left| \sum_{j=1}^{k} Z \circ T^j \right| \right\|_2$$

on the space of all $Z \in L_0^2$ with $\|Z\|_{M^+} < \infty$.

Theorem 1. *Assume that $(X_k)_{k\in Z}$ is a stationary sequence of centered square-integrable random variables. Then*

$$\|Y_0^m\|_{M^+} \to 0 \qquad \text{as } m \to \infty \tag{7}$$

if and only if there exists a martingale with stationary increments satisfying (5). Such a martingale is unique if it exists. In particular, (7) implies (4).

As a consequence of the proof of Theorem 1, we also obtain the following result that adds a new equivalent condition to the characterizations by Dedecker, Merlevède and Volný [7] and Zhao and Woodroofe [32]. With $(Y_k^m)_{k\in Z}$ defined by (6) and the seminorm notation

$$\|Y_0^m\|_+ = \limsup_{n\to\infty} \frac{1}{\sqrt{n}} \left\| \sum_{j=1}^n Y_j^m \right\|_2$$

we have the following characterization.

Theorem 2. *Assume that $(X_k)_{k\in Z}$ is as in Theorem 1. Then*

$$\|Y_0^m\|_+ \to 0 \qquad \text{as } m\to\infty \tag{8}$$

if and only if there exists a stationary martingale satisfying (2). Such a martingale is unique if it exists. In particular, (8) implies (3).

Our approach is constructive. If the stationary sequence is supposed to be ergodic, then the constructed martingale differences are also ergodic and therefore the conditional theorems (3) and (4) can be easily transported to the original processes satisfying (8) and (7), respectively, with $\eta = \|D_0\|_2$.

A natural and useful question is to provide classes of stochastic processes that have a martingale decomposition with an error term satisfying (5), in other words, to provide sharp sufficient conditions for such a decomposition. Obviously, a maximal inequality is needed in order to verify this condition. We shall combine our approach with several maximal inequalities. One is due to Rio [26], formula (3.9), page 53; for related inequalities, see [23] and [9].

- For any stationary process with centered variables in $\mathbb{L}_2$,

$$\mathbb{E}\Big(\max_{1\le i\le n} S_i^2\Big) \le 8n\mathbb{E}(X_0^2) + 16\sum_{k=2}^n \mathbb{E}|X_0\mathbb{E}_0(S_k - S_1)|. \tag{9}$$

Another inequality comes from [24], Proposition (2.3); see also [25], Theorem 1, for the inequality in $\mathbb{L}_p$.

- For any stationary process with centered variables in $\mathbb{L}_2$,

$$\begin{aligned}\mathbb{E}\Big(\max_{1\le i\le n} S_i^2\Big) &\le n\left(2\|X_0\|_2 + 3\sum_{j=0}^{r-1}\frac{\|\mathbb{E}_0(S_{2^j})\|_2}{2^{j/2}}\right)^2 \\ &\le n\left(2\|X_0\|_2 + 80\sum_{j=1}^{n}\frac{\|\mathbb{E}_0(S_j)\|_2}{j^{3/2}}\right)^2,\end{aligned} \tag{10}$$

where $2^{r-1} < n \leq 2^r$.

The following maximal inequality is a particular case of Dedecker and Merlevède [6], Proposition 6; see [34], Theorem 1, for the inequality in $\mathbb{L}_p$.

- For any stationary process with centered variables in $\mathbb{L}_2$ such that $\mathbb{E}(X_0|\mathcal{F}_{-\infty}) = 0$ almost surely, we have

$$\mathbb{E}\Big(\max_{1\leq i\leq n} S_i^2\Big) \leq 4n\left(\sum_{i=0}^{\infty} \|\mathbb{E}_{-i}(X_0) - \mathbb{E}_{-i-1}(X_0)\|_2\right)^2. \tag{11}$$

Another inequality we use for additive functionals of stationary reversible Markov chains is a consequence of Wu [33], Corollary 2.7 and relation (2.5) in the same paper (note that there is a typographical error in this relation, namely, a square should be added to the norm); see also [28]:

- Assume $(\xi_n)_{n\in\mathbb{Z}}$ is a stationary, reversible Markov chain and $X_n = f(\xi_n)$ with $f \in \mathbb{L}_0^2(\pi)$. Then, for every $n \geq 1$,

$$\mathbb{E}\Big(\max_{1\leq i\leq n} S_i^2\Big) \leq (24n+3)\sum_{n=0}^{\infty} \mathbb{E}(X_0 X_n), \tag{12}$$

  provided the series on the right-hand side is convergent.

This inequality, originally stated for the ergodic case, extends without changes to the general case.

By combining the martingale decomposition in Theorem 1 with these maximal inequalities, we point out various classes of stochastic processes for which a conditional functional limit theorem holds. These include mixing processes and classes of Markov chains.

2. Proof of Theorem 1

The proof of this theorem has several steps.

Step 1. *Construction of the approximating martingale.*

The construction of the martingale decomposition is based on averages. It was introduced by Wu and Woodroofe [35] (see their definition (6) on page 1677) and further developed in [32], extending the construction in [12] and [17]; see also [3], Theorem 8.1, and [18]. We give the martingale construction here for completeness.

We introduce a parameter $m \geq 1$ (kept fixed for the moment) and define the following stationary sequence of random variables:

$$\theta_0^m = \frac{1}{m}\sum_{i=1}^{m} \mathbb{E}_0(S_i), \theta_k^m = \theta_0^m \circ T^k.$$

Set

$$D_k^m = \theta_{k+1}^m - \mathbb{E}_k(\theta_{k+1}^m), \qquad M_n^m = \sum_{k=0}^{n-1} D_k^m. \tag{13}$$

Then $(D_k^m)_{k\in\mathbb{Z}}$ is a stationary martingale difference sequence and $(M_n^m)_{n\geq 0}$ is a martingale. Thus, we have

$$X_k = D_k^m + \theta_k^m - \theta_{k+1}^m + \frac{1}{m}\mathbb{E}_k(S_{k+m+1} - S_{k+1})$$

and therefore

$$\begin{aligned} S_k &= M_k^m + \theta_0^m - \theta_k^m + \sum_{j=1}^{k} \frac{1}{m}\mathbb{E}_{j-1}(S_{j+m} - S_j) \\ &= M_k^m + \theta_0^m - \theta_k^m + \overline{R}_k^m, \end{aligned} \tag{14}$$

where we have made use of the notation

$$\overline{R}_k^m = \sum_{j=1}^{k} \frac{1}{m}\mathbb{E}_{j-1}(S_{j+m} - S_j).$$

Observe that

$$\overline{R}_k^m = \sum_{j=0}^{k-1} Y_j^m. \tag{15}$$

With the notation

$$R_k^m = \theta_0^m - \theta_k^m + \overline{R}_k^m, \tag{16}$$

we have

$$S_k = M_k^m + R_k^m. \tag{17}$$

Step 2. *Sufficiency.*

We show that $\|Y_0^m\|_{M^+} \to 0$ as $m \to \infty$ is sufficient for (5).

The starting point is the construction of the martingale differences, as in (13). By the martingale property and (17), for all positive integers m' and m'', we have

$$\|D_0^{m'} - D_0^{m''}\|_2 = \frac{1}{\sqrt{n}}\|M_n^{m'} - M_n^{m''}\|_2 = \frac{1}{\sqrt{n}}\|R_n^{m'} - R_n^{m''}\|_2.$$

We now let $n \to \infty$. By relation (16) and stationarity, it follows that

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{\sqrt{n}}\|R_n^{m'} - R_n^{m''}\|_2 &= \limsup_{n\to\infty} \frac{1}{\sqrt{n}}\|\overline{R}_n^{m'} - \overline{R}_n^{m''}\|_2 \\ &\leq \limsup_{n\to\infty} \frac{1}{\sqrt{n}}(\|\overline{R}_n^{m'}\|_2 + \|\overline{R}_n^{m''}\|_2). \end{aligned}$$

By (7), the limit when m' and m'' both tend to ∞ is then 0, giving that (D_0^m) is Cauchy in $\mathbb{L}_2$ and therefore convergent. Denote its limit by D_0. Then $M_n = \sum_{k=0}^{n-1} D_k$ is a martingale with the desired properties. To see this, we start from the decomposition in relation (14) and obtain

$$|S_k - M_k| \le |M_k^m - M_k| + |\theta_k^m - \theta_0^m| + |\overline{R}_k^m|.$$

Then

$$\begin{aligned}\frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n} |S_k - M_k|\Big\|_2 &\le \frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n} |M_k^m - M_k|\Big\|_2 \\ &\quad + \frac{1}{\sqrt{n}}\|\theta_0^m\|_2 + \frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n} |\theta_k^m|\Big\|_2 + \frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n} |\overline{R}_k^m|\Big\|_2.\end{aligned}$$

By Doob's maximal inequality for martingales and by stationarity, we conclude that

$$\frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n} |M_k^m - M_k|\Big\|_2 \le \|D_0^m - D_0\|_2.$$

For m fixed, since $(\theta_k^m)_{k\in Z}$ is a stationary sequence of square-integrable random variables, for any $A>0$, we have

$$\begin{aligned}\frac{1}{n}\mathbb{E}\Big[\max_{1\le k\le n} |\theta_k^m|^2\Big] &\le \frac{A^2}{n} + \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}[|\theta_k^m|^2 I(|\theta_k^m| > A)] \\ &= \frac{A^2}{n} + \mathbb{E}[|\theta_0^m|^2 I(|\theta_0^m| > A)]\end{aligned}$$

and then, clearly,

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\Big[\max_{1\le k\le n} |\theta_k^m|^2\Big] = 0. \tag{18}$$

Then, taking into account (15), we easily obtain

$$\limsup_{n\to\infty} \frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n} |S_k - M_k|\Big\|_2 \le \|D_0^m - D_0\|_2 + \|Y_0^m\|_{M^+}$$

and the result follows by letting $m\to\infty$, from the fact that $D_0^m \to D_0$ in $\mathbb{L}_2$. It is easy to see that the martingale is unique.

Step 3. *Necessity.*

Assume that the martingale approximation (5) holds. With the notation $R_n = S_n - M_n$, we then have

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n} |R_k|\Big\|_2 = 0.$$

In particular, this approximation implies that

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \max_{1\le k\le n} \|\mathbb{E}(S_k|\mathcal{F}_0)\|_2 = 0. \tag{19}$$

From

$$\|\overline{R}_n^n\|_2 \le \|\mathbb{E}(S_n|\mathcal{F}_0)\|_2,$$

we deduce that

$$\|R_n^n\|_2 = \|\theta_0^n - \theta_n^n + \overline{R}_n^n\|_2 \le 2\|\theta_0^n\|_2 + \|\overline{R}_n^n\|_2 \le 3 \max_{1\le k\le n} \|\mathbb{E}(S_k|\mathcal{F}_0)\|_2,$$

whence, by (19), it follows that

$$\lim_{n\to\infty} \frac{\|R_n^n\|_2}{\sqrt{n}} = 0.$$

As a consequence, we obtain

$$\mathbb{E}(D_0^n - D_0)^2 = \frac{\mathbb{E}(M_n^n - M_n)^2}{n} = \frac{\mathbb{E}(R_n^n - R_n)^2}{n} \to 0 \qquad \text{as } n \to \infty.$$

This shows that $D_0^n \to D_0$ in $\mathbb{L}_2$. By the triangle inequality, followed by Doob's inequality, for any positive integer m, we have

$$\begin{aligned}\frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n}|R_k^m|\Big\|_2 &\le \frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n}|R_k|\Big\|_2 + \frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n}|M_k^m - M_k|\Big\|_2\\ &\le \frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n}|R_k|\Big\|_2 + \|D_0^m - D_0\|.\end{aligned}$$

Now, letting $n\to\infty$ followed by $m\to\infty$, we obtain

$$\lim_{m\to\infty}\limsup_{n\to\infty}\frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n}|R_k^m|\Big\|_2 = 0. \tag{20}$$

Now, observe that by (16), $R_n^m - \overline{R}_n^m = \theta_0^m - \theta_n^m$. Then, for every fixed m, by (18), we have

$$\frac{1}{\sqrt{n}}\Big\|\max_{1\le k\le n}|\theta_0^m - \theta_k^m|\Big\|_2 \underset{n\to\infty}{\to} 0.$$

Thus, we conclude from (20) that

$$\lim_{m\to\infty}\|Y_0^m\|_{M^+} = 0$$

and the necessity follows.

3. Applications

3.1. Applications using projective criteria

The first application involves the class of variables satisfying the Maxwell–Woodroofe condition [19].

Proposition 3. *Assume that*

$$\Delta(X_0) = \sum_{k=1}^{\infty} \frac{\|\mathbb{E}_0(S_k)\|_2}{k^{3/2}} < \infty. \tag{21}$$

The martingale approximation (5) then holds.

Proof. In order to verify condition (7) of Theorem 1, we apply inequality (10) to the stationary sequence $(Y_k^m)_{k\in\mathbb{Z}}$ defined by (6). Then

$$\left\| \max_{1\le j\le n} \left| \sum_{k=0}^{j-1} Y_k^m \right| \right\|_2 \le n^{1/2}(2\|Y_0^m\|_2 + 80\Delta(Y_0^m)).$$

First, note that by Peligrad and Utev [24], Proposition 2.5, we know that condition (21) implies that $\|Y_0^m\|_2 \to 0$. We complete the proof by showing that

$$\Delta(Y_0^m) \underset{m\to\infty}{\longrightarrow} 0.$$

Since $\|Y_0^m\|_2 \to 0$, by the triangle inequality and stationarity, every term of the series on the right-hand side of the equality

$$\Delta(Y_0^m) = \sum_{k=1}^{\infty} \frac{1}{k^{3/2}} \|\mathbb{E}_0(Y_0^m + \cdots + Y_{k-1}^m)\|_2$$

tends to 0 as $m \to \infty$. Furthermore, because

$$\begin{aligned} \|\mathbb{E}_0(Y_0^m + \cdots + Y_{k-1}^m)\|_2 &= \left\| \mathbb{E}_0 \left(\frac{1}{m} \sum_{l=1}^{m} \sum_{i=0}^{k-1} \mathbb{E}_i(X_{i+l}) \right) \right\|_2 \\ &\le \|\mathbb{E}_0(X_0 + \cdots + X_{k-1})\|_2, \end{aligned}$$

each term in $\Delta(Y_0^m)$ is dominated by the corresponding term in $\Delta(X_0)$, the latter being independent of m. The result follows from the above considerations, along with the Lebesgue dominated convergence theorem for the counting measure. □

For the sake of applications, we give the following corollary.

Corollary 4. *Assume that*

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \|\mathbb{E}_0(X_n)\|_2 < \infty. \tag{22}$$

The martingale representation (5) then holds.

The fact that (22) implies (21) was observed in Maxwell and Woodroofe [19].

We shall now combine Theorem 1 with Rio's maximal inequality (9) to obtain the following proposition.

Proposition 5. *Assume that for any $j \geq 0$,*

$$\Gamma_j = \sum_{k \geq j} \|X_j \mathbb{E}_0(X_k)\|_1 < \infty \quad \textit{and} \quad \frac{1}{m} \sum_{j=0}^{m-1} \Gamma_j \to 0 \qquad \textit{as } m \to \infty. \tag{23}$$

The martingale representation (5) then holds.

Proof. In order to verify condition (7), we now apply the maximal inequality (9) to $(Y_k^m)_{k \geq 1}$ defined by (6). We conclude that for $n \geq m$,

$$\begin{aligned} \left\| \max_{1 \leq j \leq n} \left| \sum_{k=0}^{j-1} Y_k^m \right| \right\|_2^2 &\leq 8n \|Y_0^m\|_2^2 + 16 \sum_{j=1}^{n-1} \|Y_0^m \mathbb{E}_0(Y_1^m + \cdots + Y_j^m)\|_1 \\ &\leq 8n(12m+1)\|Y_0^m\|_2^2 + 16 \sum_{j=m+1}^{n-1} \|Y_0^m \mathbb{E}_0(Y_{m+1}^m + \cdots + Y_j^m)\|_1, \end{aligned}$$

where, in the last sum, we have implemented a decomposition into two terms to deal with overlapping blocks. So, for an absolute constant C,

$$\frac{1}{n} \left\| \max_{1 \leq j \leq n} \left| \sum_{k=0}^{j-1} Y_k^m \right| \right\|_2^2 \leq C \left(\frac{\|\mathbb{E}_0(S_m)\|_2^2}{m} + \frac{1}{n} \sum_{l=m+1}^{n-1} \|Y_0^m \mathbb{E}_0(Y_{m+1}^m + \cdots + Y_l^m)\|_1 \right).$$

Since, for any $l > m$,

$$\begin{aligned} \|Y_0^m \mathbb{E}_0(Y_{m+1}^m + \cdots + Y_l^m)\|_1 &\leq \frac{1}{m} \sum_{j=1}^{m} \sup_{i > m} \|(\mathbb{E}_0(X_j))\mathbb{E}_0(X_i + \cdots + X_{i+l})\|_1 \\ &\leq \frac{1}{m} \sum_{j=1}^{m} \sum_{k \geq m} \|\mathbb{E}_0(X_j)\mathbb{E}_0(X_k)\|_1 \end{aligned}$$

and also

$$\|\mathbb{E}_0(S_m)\|_2^2 \leq 2 \sum_{j=0}^{m-1} \sum_{k=j}^{m-1} \|\mathbb{E}_0(X_j)\mathbb{E}_0(X_k)\|_1,$$

we then obtain, by the properties of conditional expectations, that for a certain absolute constant C',

$$\frac{1}{n} \left\| \max_{1 \leq j \leq n} \left| \sum_{k=0}^{j-1} Y_k^m \right| \right\|_2^2 \leq \frac{C'}{m} \sum_{j=0}^{m} \sum_{k \geq j} \|X_j \mathbb{E}_0(X_k)\|_1$$

and the result follows from condition (23), by first letting $n \to \infty$, followed by $m \to \infty$. □

The projective criteria in the next proposition were studied in [11, 13, 16], among others.

Proposition 6. *Assume*

$$\mathbb{E}(X_0|\mathcal{F}_{-\infty}) = 0 \quad \textit{almost surely and} \quad \sum_{i=1}^{\infty} \|\mathbb{E}_{-i}(X_0) - \mathbb{E}_{-i-1}(X_0)\|_2 < \infty. \tag{24}$$

The martingale approximation (5) then holds.

Proof. The validity of this proposition easily follows by verifying condition (7) via maximal inequality (11) applied to $(Y_k^m)_{k\geq 1}$ defined by (6). Indeed, by (11), the triangle inequality and stationarity, we have

$$\begin{aligned}\frac{1}{\sqrt{n}}\left\|\max_{1\leq j\leq n}\left|\sum_{k=0}^{j-1} Y_k^m\right|\right\|_2 &\leq 2\sum_{i=0}^{\infty} \|\mathbb{E}_{-i}(Y_0^m) - \mathbb{E}_{-i-1}(Y_0^m)\|_2 \\ &\leq \frac{2}{m}\sum_{i=0}^{\infty}\sum_{k=1}^{m} \|\mathbb{E}_{-i}(X_k) - \mathbb{E}_{-i-1}(X_k)\|_2.\end{aligned}$$

Now, by stationarity, change of order of summation and change of variable,

$$\frac{1}{\sqrt{n}}\left\|\max_{1\leq j\leq n}\left|\sum_{k=0}^{j-1} Y_k^m\right|\right\|_2 \leq \frac{2}{m}\sum_{k=1}^{m}\sum_{j=k}^{\infty} \|\mathbb{E}_{-j}(X_0) - \mathbb{E}_{-j-1}(X_0)\|_2.$$

To verify condition (7), we let $n \to \infty$ followed by $m \to \infty$. Note that the term on the right-hand side of the previous inequality tends to 0 as $m \to \infty$, by (24). □

3.2. Application to mixing sequences

The results in the previous section can be immediately applied to mixing sequences, leading to the sharpest possible results and providing additional information about the structures of these processes. Examples include various classes of Markov chains and Gaussian processes.

We shall introduce the following mixing coefficients: for any two σ-algebras $\mathcal{A}$ and $\mathcal{B}$, define the strong mixing coefficient $\alpha(\mathcal{A}, \mathcal{B})$,

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|; A \in \mathcal{A}, B \in \mathcal{B}\},$$

and the ρ-mixing coefficient, known also as the maximal coefficient of correlation $\rho(\mathcal{A}, \mathcal{B})$,

$$\rho(\mathcal{A}, \mathcal{B}) = \sup\{\mathrm{Cov}(X, Y)/\|X\|_2\|Y\|_2 : X \in \mathbb{L}_2(\mathcal{A}), Y \in \mathbb{L}_2(\mathcal{B})\}.$$

For the stationary sequence of random variables $(X_k)_{k\in\mathbb{Z}}$, we also define $\mathcal{F}_m^n$, the σ-field generated by X_i with indices $m \le i \le n$. $\mathcal{F}^n$ denotes the σ-field generated by X_i with indices $i \ge n$ and $\mathcal{F}_m$ denotes the σ-field generated by X_i with indices $i \le m$. The sequences of coefficients $\alpha(n)$ and $\rho(n)$ are then defined by

$$\alpha(n) = \alpha(\mathcal{F}_0, \mathcal{F}_n^n) \quad \text{and} \quad \rho(n) = \rho(\mathcal{F}_0, \mathcal{F}^n).$$

Equivalently (see [2], Chapter 4),

$$\rho(n) = \sup\{\|\mathbb{E}(Y|\mathcal{F}_0)\|_2/\|Y\|_2 : Y \in \mathbb{L}_2(\mathcal{F}^n),\ \mathbb{E}(Y) = 0\}.$$

Finally, we say that the stationary sequence is strongly mixing if $\alpha(n) \to 0$ as $n \to \infty$ and ρ-mixing if $\rho(n) \to 0$ as $n \to \infty$.

An interesting application of Proposition 3 is to ρ-mixing sequences. It is well known that the central limit theorem and its invariance principle hold for stationary centered sequences with finite second moments under the assumption

$$\sum_{k=1}^{\infty} \rho(2^k) < \infty, \tag{25}$$

where $\rho(n) = \rho(\mathcal{F}_0, \mathcal{F}^n)$. Let us recall that the central limit theorem is due to [14], while the invariance principle is found in [22, 29–31]. The fact that condition (25) is sharp in this context is due to [2], Volume 1, page 367, and Volume 3, Theorem 34.13. Bradley's example shows that if (25) fails, then $S_n/\|S_n\|_2$ might have non-degenerate non-normal distributions as weak limit points.

As a corollary of Proposition 3, we obtain the conditional invariance principle for ρ-mixing sequences.

Proposition 7. *Assume $\sum_{k=1}^{\infty} \rho(2^k) < \infty$. The martingale representation (5) then holds.*

Proof. As in [21], for a positive constant C, we have

$$\sum_{r=0}^{\infty} \frac{\|\mathbb{E}(S_{2^r}|\mathcal{F}_0)\|_2}{2^{r/2}} \le C\|X_0\|_2 \sum_{j=0}^{\infty} \rho(2^j).$$

□

To obtain sharp results for strongly mixing sequences, we shall use Proposition 5.

According to Doukhan, Massart and Rio [10], a condition that is optimal for CLT or the invariance principle for strongly mixing sequences is

$$\sum_{k\ge 1} \mathbb{E}X_0^2 I(|X_0| \ge Q_{|X_0|}(2\alpha_k)) < \infty, \tag{26}$$

where $Q_{|X_0|}$ denotes the cadlag inverse of the function $t \to P(|X_0| > t)$. Also under this condition, we add the additional information given by Theorem 1.

Proposition 8. *Assume that condition (26) is satisfied. The martingale representation (5) then holds.*

Proof. We shall just verify the condition of Proposition 5. Note that on the set $[0,\ P(|Y|>0)]$, the function $H_Y : x \to \int_0^x Q_Y(u)\,\mathrm{d}u$ is an absolutely continuous and increasing function with values in $[0,\ E|Y|]$. Denote by G_Y the inverse of H_Y. With this notation, by Merlevède and Peligrad [20], relation (4.84), we have

$$\|X_j \mathbb{E}(X_k|\mathcal{F}_0)\|_1 \le 3 \int_0^{\|\mathbb{E}(X_k|\mathcal{F}_0)\|_1} Q_{|X_0|} \circ G(u)\,\mathrm{d}u$$

and we then majorize the right-hand side in the previous inequality by Dedecker and Doukhan [5], Proposition 1, to obtain

$$\|X_j \mathbb{E}(X_k|\mathcal{F}_0)\|_1 \le 6 \int_0^{2\alpha(k)} Q^2_{|X_0|}\,\mathrm{d}u.$$

Therefore,

$$\begin{aligned}\sum_{k\ge j} \|X_j \mathbb{E}_0(X_k)\|_1 &\le 6 \sum_{k\ge j} \int_0^{2\alpha(k)} Q^2_{|X_0|}\,\mathrm{d}u \\ &\le 6 \sum_{k\ge j} \mathbb{E} X_0^2 I(|X_0| \ge Q_{|X_0|}(2\alpha_k)) \to 0 \qquad \text{as } j\to\infty.\end{aligned}$$

□

Note that the coefficient $\alpha(k)$ is defined by using only one variable in the future. Moreover, by the Cauchy–Schwarz inequality, condition (26) is satisfied if the variables have finite moments of order $2+\delta$ for a $\delta>0$ and

$$\sum_{k\ge 1} \alpha(k)^{\delta/(2+\delta)} < \infty.$$

An excellent source of information for classes of mixing sequences and classes of Markov chains satisfying mixing conditions is the book by Bradley [2]. Further applications can be obtained by using the coupling coefficients in [8].

3.3. Application to additive functionals of reversible Markov chains

For reversible Markov processes (i.e., $Q=Q^*$), the invariance principle under an optimal condition is known since Kipnis and Varadhan [18]. The following is a formulation in terms of martingale approximation.

Proposition 9. *Let $(\xi_i)_{i\in\mathbb{Z}}$ be a stationary reversible Markov chain and $f \in \mathbb{L}_0^2(\pi)$ with the property*

$$\lim_{n\to\infty} \frac{\operatorname{var}(S_n)}{n} \to \sigma_f^2 < \infty. \tag{27}$$

The martingale approximation satisfying (5) then holds.

Proof. We have to verify condition (7). Denote by ρ_f the spectral measure of f corresponding to the self-adjoint operator Q on $\mathbb{L}_2(\pi)$. It is well known that the assumption (27) for $f \in \mathbb{L}_0^2$ implies that $\int_{-1}^{1}(1-t)^{-1}\rho_f(\mathrm{d}t) < \infty$ (see [18]). Define Y_0^m by (6). By the maximal inequality (12), we have

$$\frac{1}{n}\mathbb{E}\left(\max_{1\le j\le n}\left|\sum_{k=0}^{j-1} Y_k^m\right|\right)^2 \le 27\sum_{k\ge 0}\mathbb{E}(Y_0^m Y_k^m),$$

provided that the sum on the right-hand side is finite. To prove it, by using spectral calculus for the self-adjoint operator Q, we obtain

$$\sum_{k\ge 0}\mathbb{E}(Y_0^m Y_k^m) \le \frac{1}{m^2}\int_{-1}^{1}\frac{(1+t+\cdots+t^{m-1})^2}{(1-t)}\rho_f(\mathrm{d}t)$$

and therefore, for every positive integer $m > 0$,

$$\|Y_0^m\|_{M^+}^2 \le 27\int_{-1}^{1}\frac{(1+t+\cdots+t^{m-1})^2}{m^2(1-t)}\rho_f(\mathrm{d}t).$$

Since $\int_{-1}^{1}(1-t)^{-1}\rho_f(\mathrm{d}t) < \infty$, the right-hand side is finite and, by the dominated convergence theorem,

$$\lim_{m\to\infty}\|Y_0^m\|_{M^+}^2 = 0. \qquad \square$$

Similar results are expected to hold for other classes of stationary and ergodic Markov chains when Q is not necessarily self-adjoint, but instead satisfies a quasi-symmetry or strong sector condition, or is symmetrized. See [33] and [28] for these related processes.

3.4. Application to additive functionals of normal Markov chains

For additive functionals of normal Markov chains ($QQ^* = Q^*Q$), the central limit theorem below is a result of Gordin and Lifshitz [17]. As an application of Theorem 2, we give an alternative proof.

Let ρ_f be the spectral measure on the closed unit disk $D \subset \mathbb{C}$ corresponding to the function $f \in \mathbb{L}_0^2(\pi)$.

Proposition 10. *Let* $(\xi_i)_{i\in\mathbb{Z}}$ *be a stationary normal Markov chain and a function* $f \in \mathbb{L}_0^2(\pi)$, *satisfying the condition*

$$\int_D \frac{1}{|1-z|}\rho_f(\mathrm{d}z) < \infty. \tag{28}$$

The martingale approximation (2) then holds.

Proof. According to Theorem 2, we have to verify condition (8). By using spectral calculus as in [3], Chapter 4, after some computations, we get

$$\limsup_{n\to\infty} \frac{1}{n}\left\|\sum_{k=0}^{n-1} Y_k^m\right\|_2^2 \leq 4\int_D \frac{|1+z+\cdots+z^{m-1}|^2}{m^2|1-z|}\rho_f(\mathrm{d}z)$$

and condition (8) is therefore satisfied by condition (28) and the dominated convergence theorem. □

Condition (28) has an interesting equivalent formulation in terms of conditional moments that is in the spirit of (and which implies) the Mawxell–Woodroofe condition (21).

Remark 11. Condition (28) is equivalent to

$$\sum_{k=1}^{\infty} \frac{\|\mathbb{E}_0(S_k)\|_2^2}{k^2} < \infty. \tag{29}$$

Condition (29) is further implied by

$$\sum_{k=1}^{\infty} \|\mathbb{E}_0(X_k)\|_2^2 < \infty. \tag{30}$$

The equivalence in the above remark can be found in [4], Lemma 2.1. The fact that (30) implies (29) is easily established, much like the proof that (22) implies (21).

Acknowledgements

Mikhail Gordin was supported in part by a Charles Phelps Taft Memorial Fund grant and RFBR Grant 10-01-00242_a. Magda Peligrad was supported in part by a Charles Phelps Taft Memorial Fund grant and NSA Grant H98230-09-1-0005. The authors are grateful to the referees for carefully reading the paper and for numerous suggestions that improved the presentation of the paper.

References

[1] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. MR0233396

[2] Bradley, R.C. (2007). *Introduction to Strong Mixing Conditions*, Vols 1–3. Heber City, UT: Kendrick Press. MR2325294

[3] Borodin, A.N. and Ibragimov, I.A. (1994). Limit theorems for functionals of random walks. *Trudy Mat. Inst. Steklov.* **195** 286. Transl. in English: *Proc. Steklov Inst. Math.* **195** (1995). MR1368394

[4] Cuny, C. (2009). Pointwise ergodic theorems with rate and application to limit theorems for stationary processes. Available at arXiv:0904.0185v1. MR2542901

[5] Dedecker, J. and Doukhan, P. (2003). A new covariance inequality and applications. *Stochastic Process. Appl.* **106** 63–80. MR1983043

[6] Dedecker, J. and Merlevède, F. (2002). Necessary and sufficient conditions for the conditional central limit theorem. *Ann. Probab.* **30** 1044–1081. MR1920101

[7] Dedecker, J., Merlevède, F. and Volný, D. (2007). On the weak invariance principle for non-adapted stationary sequences under projective criteria. *J. Theoret. Probab.* **20** 971–1004. MR2359065

[8] Dedecker, J. and Prieur, C. (2005). New dependence coefficients. Examples and applications to statistics. *Probab. Theory Related Fields* **132** 203–236. MR2199291

[9] Dedecker, J. and Rio, E. (2000). On the functional central limit theorem for stationary processes. *Ann. Inst. H. Poincaré Probab. Statist.* **36** 1–34. MR1743095

[10] Doukhan, P., Massart, P. and Rio, E. (1994). The functional central limit theorem for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.* **30** 63–82. MR1262892

[11] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*. New York: Academic Press. MR0624435

[12] Heyde, C.C. (1974). On the central limit theorem for stationary processes. *Z. Wahrsch. Verw. Gebiete.* **30** 315–320. MR0372955

[13] Hannan, E.J. (1979). The central limit theorem for time series regression. *Stochastic Process. Appl.* **9** 281–289. MR0562049

[14] Ibragimov, I.A. (1975). A note on the central limit theorem for dependent variables. *Theory Probab. Appl.* **20** 135–140. MR0362448

[15] Gordin, M.I. (1969). The central limit theorem for stationary processes. *Soviet. Math. Dokl.* **10** 1174–1176. MR0251785

[16] Gordin, M.I. (2004). A remark on the martingale method for proving the central limit theorem for stationary sequences. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov (POMI)* **311**. *Veroyatn. i Stat.* **7** 124–132, 299–300. Transl.: *J. Math. Sci. (N.Y.)* **133** (2006) 1277–1281. MR2092203

[17] Gordin, M.I. and Lifshitz, B. (1981). A remark about a Markov process with normal transition operator. In *Third Vilnius Conf. Probab. Stat.* **1** 147–148. Vilnius: Akad. Nauk Litovsk.

[18] Kipnis, C. and Varadhan, S.R.S. (1986). Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104** 1–19. MR0834478

[19] Maxwell, M. and Woodroofe, M. (2000). Central limit theorems for additive functionals of Markov chains. *Ann. Probab.* **28** 713–724. MR1782272

[20] Merlevède, F. and Peligrad, M. (2006). On the weak invariance principle for stationary sequences under projective criteria. *J. Theoret. Probab.* **19** 647–689. MR2280514

[21] Merlevède, F., Peligrad, M. and Utev, S. (2006). Recent advances in invariance principles for stationary sequences. *Probab. Surv.* **3** 1–36. MR2206313

[22] Peligrad, M. (1982). Invariance principle for mixing sequences of random variables. *Ann. Probab.* **10** 968–981. MR0672297

[23] Peligrad, M. (1999). Convergence of stopped sums of weakly dependent random variables. *Electron. J. Probab.* **4** 1–13. MR1692676

[24] Peligrad, M. and Utev, S. (2005). A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.* **33** 798–815. MR2123210

[25] Peligrad, M., Utev, S. and Wu, W.B. (2007). A maximal L_p-inequality for stationary sequences and its applications. *Proc. Amer. Math. Soc.* **135** 541–550. MR2255301

[26] Rio, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants*, Mathématiques & Applications **31**. Berlin: Springer. MR2117923

[27] Rootzén, H. (1976). Fluctuations of sequences which converge in distribution. *Ann. Probab.* **4** 456–463. MR0410865

[28] Sethuraman, S., Varadhan, S.R.S. and Yau, H.T. (2000). Difusive limit of a tagged particle in asymmetric simple exclusion processes. *Comm. Pure Appl. Math.* **53** 972–1006. MR1755948

[29] Shao, Q. (1989). On the invariance principle for stationary ρ-mixing sequences of random variables. *Chinese Ann. Math. Ser. B* **10B** 427–433. MR1038376

[30] Utev, S.A. (1989). Sums of random variables with φ-mixing. *Trudy Inst. Mat.* **1** 78–100. MR1037250

[31] Utev, S.A. (1991). Sums of random variables with φ-mixing. *Siberian Adv. Math.* **1** 124–155. MR1128381

[32] Zhao, O. and Woodroofe, M. (2008). On martingale approximations. *Ann. Appl. Probab.* **18** 1831–1847. MR2462550

[33] Wu, L. (1999). Forward–backward martingale decomposition and compactness results for additive functionals of stationary ergodic Markov processes. *Ann. Inst. H. Poincaré Probab. Statist.* **35** 121–141. MR1678517

[34] Wu, W.B. (2007). Strong invariance principles for dependent random variables. *Ann. Probab.* **35** 2294–2320. MR2353389

[35] Wu, W.B. and Woodroofe, M. (2004). Martingale approximations for sums of stationary processes. *Ann. Probab.* **32** 1674–1690. MR2060314

Received July 2009 and revised February 2010